PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, C12P 19/34 — A1

(11) International Publication Number: WO 97/20076

(43) International Publication Date: 5 June 1997 (05.06.97)

(21) International Application Number: PCT/US96/19030

(22) International Filing Date: 27 November 1996 (27.11.96)

(30) Priority Data:
60/007,694 29 November 1995 (29.11.95) US

(71) Applicant *(for all designated States except US)*: NEXSTAR PHARMACEUTICALS, INC. [US/US]; Suite 200, 2860 Wilderness Place, Boulder, CO 80301 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: SCHMIDT, Paul [US/US]; 1730 Euclid Avenue, San Marino, CA 91108 (US).

(74) Agents: SWANSON, Barry, J. et al.; Suite 200, 8400 E. Prentice Avenue, Englewood, CO 80111 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHODS FOR OPTIMIZING MULTICOMPONENT FORMULATIONS

(57) Abstract

Methods for optimizing multicomponent formulations are described. The invention comprises the simultaneous analysis of a large number of mixtures to determine the formulation(s) with optimal properties. A unique identification marker, capable of being accurately analyzed with higher sensitivity than the formulation itself, is associated with each mixture. The optimal formulation is identified by analysis of the marker, which is specific for a particular set and ratio of ingredients.

BNSDOCID: <WO_____9720076A1_I_>

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

BNSDOCID: <WO_____9720076A1_I_>

METHODS FOR OPTIMIZING MULTICOMPONENT FORMULATIONS

FIELD OF THE INVENTION

The present invention relates to methods for identifying optimal formulations of multicomponent systems. An optimal formulation is a particular combination of two or more component materials mixed in a specific ratio such that the formulation optimally possesses a desired set of properties.

BACKGROUND

Formulations comprising two or more materials are commonly encountered in the chemical and pharmaceutical fields and find use in a wide range of industries, including, for example, the biotechnological, medical, petroleum and paint industries. These multicomponent formulations often require optimization, namely, that a particular set of ingredients be combined in a preferred specific ratio in order to achieve or maximize a desired set of properties.

One particularly important example of a multicomponent formulation in the pharmaceutical field is that of drug carriers, which may, for example, be composed of polymers, proteins or lipids associated to form discrete assemblies (e.g., liposomes, nanoparticles, emulsions or colloids). Where targeted drug delivery is the goal, it is typically desired that the drug carriers be designed to preferentially target certain tissues and cells (e.g., tumors or cancerous organs) and to avoid others (e.g., normal tissues and cells). Thus, it is often necessary to determine the particular carrier formulation which reaches a target with the greatest efficiency and selectivity. Such a goal can be accomplished by formulation optimization. More generally, formulation optimization is typically required whenever it is desirable to achieve or maximize a specific set of properties for a multicomponent system.

The ability to obtain an optimal set of properties for a

multicomponent formulation is often highly dependent on the chemical composition and processing methods used to make the formulation. For example, the identity and relative amounts (i.e., ratios) of ingredients can dramatically affect the properties (e.g., chemical, physical or electrical) of the resultant formulation. For these reasons, when dealing with multicomponent formulations, it is usually necessary to evaluate a very large number of candidate formulations to find a formulation which possesses a desired set of properties.

Presently, formulation optimization is often accomplished by individually and sequentially testing particular combinations and ratios of component materials by trial and error. Unfortunately, this type of sequential empirical evaluation of candidate formulations is time-consuming and inefficient, especially where several components are combined in various ratios. Determining an optimal formulation may require hundreds or even thousands (or more) of experiments, even where a skilled worker dismisses before testing a large number of candidate formulations which are unlikely to be successful. Sequential testing is especially frustrating or impractical when the experiments are lengthy or costly to perform. Moreover, individual formulation testing is usually infeasible in the case of drug formulations, especially where the drug is being tested *in vivo* (i.e., in living animals). For example, if drug targeting to a tumor site within an animal subject is a goal, evaluation of each candidate formulation would typically require excision of a tumor from a separate animal in order to determine whether the particular formulation successfully reached the tumor site in an effective amount. For reasons such as these, individual testing of candidate formulations, where evaluation of a large number of compositions is contemplated, is often impracticable. Thus, it is desired to be able to evaluate a plurality of multicomponent formulations simultaneously, thereby avoiding the need to individually

BNSDOCID: <WO_____9720076A1_I_>

and/or sequentially test candidate formulations.

Unfortunately, standard techniques for simultaneous analysis of a large number of formulations are frequently inadequate because it is often impossible to accurately determine the identity of a formulation possessing a desired set of properties (i.e., an optimal formulation) since such a determination normally requires extraction or separation from a complex mixture. This is especially true where only very small quantities of an optimal formulation are recoverable, since standard chemical analysis may lack either adequate sensitivity or accuracy, or both. Lack of analytical sensitivity or accuracy is particularly problematic where only a small initial amount of the candidate formulation is available or where the component ratios are varied in small increments between individual candidate formulations. Moreover, chemical analysis will likely be unable to differentiate or discriminate between similar (but distinct) formulations. In addition, in a case where a formulation contains naturally occurring lipids or other biomolecular components, such components are probably indistinguishable from the biological milieu *in vivo*.

Determination of an optimal formulation is also particularly difficult where the composition is degraded or changed prior to expressing the desired set of properties (e.g., in the course of reaching a target site *in vivo*). In this case, chemical analysis of the formulation which reaches the target site would probably yield an inaccurate and meaningless result; analysis, if possible, would reflect the degraded or changed composition of the formulation at its final site, rather than the initial composition of the formulation. For example, testing may be performed on one thousand different initial candidate formulations, wherein two particular formulations differ only with respect to one ingredient: Formulation 1 comprises 10% ingredient X and Formulation 2 comprises 5% ingredient X. Assume that an initial composition containing 10% X is essential for target specificity

and, consequently, that of the 1,000 different formulations injected, only Formulation 1 reaches the desired target (i.e., Formulation 1 is the optimal formulation). Assume also that in the course of reaching the target site, Formulation 1 is either stripped of half of its amount of ingredient X or that half of its ingredient X is metabolized or degraded. In this scenario, chemical analysis of the formulation reaching the target site would incorrectly reveal the optimal formulation as one which contains approximately 5% ingredient X. Therefore, in this example, although Formulation 1 is the true optimal formulation, chemical analysis would inaccurately identify Formulation 2 as the optimal formulation. A method capable of identifying the initial composition of an optimal formulation is needed where the formulation may undergo a physical or chemical change prior to being evaluated for expression of a given set of properties.

Accordingly, it has been a desideratum to provide a method to determine optimal formulations of multicomponent systems whereby a plurality of candidate formulations is evaluated simultaneously with respect to expression of a desired set of properties.

SUMMARY OF THE INVENTION

The present invention relates to methods for optimizing multicomponent formulations, including, for example, drug carriers. Multicomponent formulations for use in the present invention preferably comprise discrete assemblies such as liposomes, emulsions, nanoparticles or colloids, which remain substantially discrete during the optimization process.

Formulation optimization may be achieved by simultaneously testing a large number of candidate formulations and evaluating them with respect to certain properties. Optimal formulations are separated from other candidate formulations by the presence of a desired set of properties, such

as targeting to tumor tissue or a site of infection, or a favorable physical distribution (e.g., a formulation which is in the form of a clear solution as opposed to a precipitate or a formulation which has a particular optical absorbance). An optimal formulation is identified as to its components by an associated marker which contains a signal sequence, or other chemical information, that is uniquely correlated with the initial composition of the optimal formulation.

The present invention also permits, where there exist multiple optimal formulations, comparative analysis of the quantities and effectiveness of each optimal formulation within the set of optimal formulations. Determination of the quantity of marker associated with each of the optimal formulations (e.g., by signal intensity) can provide a histogram that is representative of the relative individual contributions of each formulation. The advantage of the present invention in this respect is the ability to discriminate between individual formulations in the resulting mixture, which is normally impossible using conventional analytical techniques that, at best, can provide only a weighted average of all formulations in the extracted material.

In practice of the present invention, each candidate formulation is associated with one or more unique markers that is specific to the particular candidate formulation and whose identity is correlated to the initial composition of the formulation. Association of a marker with the discrete assemblies of its corresponding formulation must be sufficiently strong and stable so as to prevent cross-contamination of markers between different candidate formulations and appreciable loss of marker before the identity of the optimum formulation is determined. Each marker is associated with a candidate formulation by, for example, encapsulation of the marker within or stable attachment (chemically or physically) to the discrete assemblies of the formulation. After marking, the candidate formulations are evaluated

with respect to exhibition or performance of a desired set of properties. Evaluation in the case of pharmaceutical products, particularly therapeutic or diagnostic agents, may comprise administration of the candidate formulation to a collection of live cells, either *in vitro* or *in vivo*. In the case of *in vivo* administration, the candidate formulations may be administered to an animal body intravenously or by any other appropriate administration route, including intramuscular, subcutaneous, percutaneous, oral ingestion and topical administration.

The information correlating the composition of a given formulation and its associated marker may be encoded within the unique marker itself or may be recorded and cross-referenced in a separate database. This database comprises a record of the compositional data for each candidate formulation and the identity of its corresponding marker. The database may be recorded, stored and read using conventional data storage techniques, including physical, magnetic, electronic and optical devices (e.g., magnetic tapes and optical disks). Once the identity of a marker associated with a given formulation has been determined, the compositional data for that formulation may be found by accessing the appropriate database address which cross-references the identities of the marker and its corresponding formulation. The particular format or structure of the database, however, is not important to practice of the present invention. Preparation of an appropriate database is well understood by those skilled in the art.

The marker of the present invention is preferably capable of being accurately analyzed with substantially higher sensitivity and discrimination than the components of the formulation. The increased sensitivity is derived from, for example, the ability of the marker to be amplified or copied to yield a greater amount for analysis, or by the use of markers which can be accurately identified using more sensitive and specialized analytical techniques than those existing for the formulation ingredients

themselves.

In a preferred embodiment of the invention, the marker is a synthetic nucleic acid (DNA or RNA) oligonucleotide between about 12 and 100 nucleotides in length and having a sequence corresponding to a specific composition and component ratio of a candidate formulation. A variable region (or regions) contained within the oligonucleotide marker provides a sufficient number of oligonucleotide combinations to permit their use as unique markers to the candidate formulations. For example, using the four common nucleic acid bases for synthesis of the oligonucleotide marker yields at least 4^x possible unique combinations, where x is the number of bases in the variable region(s). Thus, where x=10, there are at least 4^{10}=1,048,576 unique oligonucleotide markers available. The length of the variable region(s) may be adjusted to provide an array of at least as many oligonucleotide combinations as there are candidate formulations to be evaluated. Association of the oligonucleotide marker with a candidate formulation may be obtained by encapsulation of the marker within or attachment of the marker to the individual assemblies or particles of the formulation.

The nucleic acid marker is preferably capable of amplification by the polymerase chain reaction (PCR) or by any of various other nucleic acid multiplication schemes known in the art. In a preferred embodiment, each oligonucleotide is flanked by defined 5' and 3' fixed regions that allow primer hybridization for Klenow extension, a primer annealing site for cDNA synthesis, PCR amplification and efficient T7 RNA polymerase transcription. PCR and other amplification techniques are described in Davis *et al.* (1994) Basic Methods in Molecular Biology, 2nd edition (Appleton & Lange, Norwalk, CT), the disclosure of which is incorporated herein by reference.

In another preferred embodiment, the marker comprises an

electrophoric moiety which is associated with the discrete assemblies of a given candidate formulation. In a preferred embodiment, the electrophoric marker may comprise a variable-length alkyl chain and a functionalized polyhaloarene moiety. A set of distinctly resolvable electrophoric markers may be prepared by varying the length of the alkyl chain and the halogen substitution of the aromatic moiety. The absence or presence of an electrophoric marker essentially constitutes a binary code which can be used to record the initial composition of the candidate formulation with which it is associated.

As with the oligonucleotide markers, association of the electrophoric marker may be obtained by encapsulation of the marker within or attachment of the marker to the discrete assemblies of a candidate formulation. Alternatively, however, a component intended to be included in the formulation may itself be chemically modified to directly incorporate the electrophoric marker. In the case where the electrophoric marker is incorporated directly into the formulation, a small amount of a component suitable for inclusion in the formulation is modified (prior to preparation of the candidate formulations) to incorporate a linker to which is attached the electrophoric marker(s) or any other suitable marker. In either case, identification of the electrophoric marker may be accomplished using various techniques known to those skilled in the art for analyzing minute quantities of organic molecules, including gas chromatography (GC) and mass spectrometry (MS).

For example, identification of the electrophoric marker may be performed preferably by using capillary gas chromatography, or most preferably, because of its detection sensitivity, electron capture gas chromatography (ECGC). In the case where the electrophoric marker is incorporated directly into the formulation by chemical modification of a component, it is preferred that the linker contains a photochemically labile

bond such that the electrophoric marker is cleaved specifically upon irradiation, thereby releasing the marker element, in this case a haloarene moiety, for facile analysis.

Alternatively, a similar set of haloarenes may be used to establish markers which may be identified by mass spectrometry. In the latter case, uniquely resolvable molecular mass signals may be obtained by varying the halogen substitution of the arene moiety or by incorporating appropriate isotopic substitutions. Where mass spectrometry is utilized for detection, the techniques of matrix-assisted laser desorption/ionization (MALDI) or electrospray ionization (ESI) may be employed to generate molecular ions of the markers. One skilled in the art is familiar with the use of these techniques in conjunction with mass spectrometry.

Aromatic halocarbons similar to the electrophoric markers of the present invention have been used in other applications as molecular tags for the indexing of synthetic chemical combinatorial libraries. See, Gallop *et al.* (1994) J. Med. Chem. 37:1233-1251 and Ohlmeyer (1993) Proc. Natl. Acad. Sci. USA 90:10922-10926 .

Due to the sensitivity with which the markers of the present invention can be identified, only very small quantities of each formulation are needed for optimization. Thus, the present invention is especially useful where the formulation is expensive or, in the case of *in vivo* evaluation, where small amounts of tissue are available and/or where the formulation is toxic at high dosages.

Other features and advantages of the present invention will be apparent upon review of the detailed description of the preferred embodiment, the drawings and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

It is to be understood that the following drawings are designed for

BNSDOCID: <WO_____9720076A1_I_>

the purpose of illustration only and are not intended to define or limit the scope of the invention.

FIGURE 1 is a flowchart illustrating one of the preferred embodiments of the method of the present invention, as described in the Detailed Description below.

FIGURE 2 depicts an example of the structure of an electrophoric marker, including a photocleavable linker. A set of chromatographically resolvable markers may be generated by varying n (the length of the alkyl chain) and Ar (the halogen substitution of the arene moiety). Examples of Ar groups suitable for use in the present invention are also shown.

DETAILED DESCRIPTION

Multicomponent formulation means a preparation consisting of two or more chemical components combined in a specific ratio. A multicomponent formulation for use in the present invention preferably comprises discrete assemblies (e.g., microparticulate preparations such as liposomes, emulsions or colloids) which may be strongly associated with one or more unique identifying markers. It is further preferred that the association between each marker and its formulation is maintained throughout the process in which formulations are evaluated for a desired set of properties, and also that the assemblies are relatively stable physically and remain essentially discrete between different individual formulations (i.e., little or no chemical mixing of components between different formulations).

A candidate formulation refers to one of a plurality of multicomponent formulations that is evaluated substantially simultaneously with respect to a desired set of properties in order to determine an optimal formulation, i.e., a candidate formulation which possesses or exhibits the desired set of properties. In some cases, the desired set of properties may

comprise the lack of a given response. For example, it may be desired that a given diagnostic or imaging contrast agent not be taken up by (i.e., is excluded from) a particular organ when administered.

An optimal formulation is a candidate formulation which has been identified as possessing a desired set of properties upon evaluation. Often, there may be more than one optimal formulation, depending on the particular set of properties chosen. When multiple optimal formulations are determined to exist, the optimal formulations will likely possess similar chemical or physical attributes and constitute a family or a small number of families of related formulations. Formulation optimization refers to determination of the initial composition (e.g., component identities and their ratios) of an optimal formulation, by evaluating a plurality of candidate formulations substantially simultaneously with respect to the desired set of properties. The set of properties may consist of one or more properties (e.g., physical, chemical, electrical) which permits an optimal formulation to be distinguished, and preferably separated, from other candidate formulations. Physical separation of an optimal formulation from other candidate formulations facilitates identification of its associated marker(s) and, consequently, its initial composition. For example, where multicomponent formulations comprising microparticulates are optimized with respect to surface charge, evaluation of candidate formulations (and separation of an optimal formulation) may be accomplished by application of an electric field to a mixture of the candidate formulations.

A marker refers to a component of each formulation which contains information that is capable of identifying each individual formulation. A marker can contain a signal sequence or other chemical information that is uniquely correlated with the initial composition of the optimal formulation. Examples of suitable markers include, but are not limited to oligonucleotides, electrophoric markers and the like.

BNSDOCID: <WO_____9720076A1_I_>

Desired properties of multicomponent formulations in the present invention may include, among others, chemical, physical, electrical, and pharmacological properties. For example, in pharmaceutical or medical applications, specific targeting of a therapeutic or diagnostic agent to a particular site *in vivo* is often desirable. In the case of drug carrier formulations, liposomes (and certain other lipid-based carriers such as emulsions) are known to be physiologically compatible and effective as delivery systems for a broad range of drugs and biomolecular agents, and may be used to encapsulate (or incorporate) both aqueous-soluble and lipid-soluble materials. For example, United States Patent No. 5,019,369, issued 5/28/91 entitled "Method of Targeting Tumors in Humans," teaches the targeting of diagnostic and chemotherapeutic agents to tumors in the body of a patient by intravenous administration of the liposomally encapsulated agent. Various methods for associating or incorporating molecules into liposomes and other lipid-based carriers are well known in the art. In addition, United States Patent No. 5,435,989, issued 7/25/95 entitled "Method of Targeting a Specific Location in a Body," and United States Patent No. 5,441,745, issued 8/15/95 entitled "Method of Delivering Micellular Particles Encapsulating Chemotherapeutic Agents to Tumors in a Body," describe the use of liposomes to deliver anticancer agents preferentially to a tumor site. The disclosures of the aforementioned United States Patents are incorporated herein by reference.

According to the present invention, methods for optimizing multicomponent formulations are described. Specifically, a plurality of candidate formulations is prepared, each formulation differing with respect to chemical composition (i.e., identity and/or ratio of ingredients). The present invention is particularly advantageous over existing techniques (e.g., sequential trial and error evaluation) where the number of candidate formulations exceeds about 10. The candidate formulations may be

prepared by random mixing of ingredients (which additionally requires chemical analysis prior to testing, in order to determine the composition of each formulation) or, preferably, by deliberate choice of formulation compositions. The appropriate volume for each individual candidate formulation may be readily determined by one skilled in the art, so as to provide a sufficient concentration of marker for evaluation and for identification by the particular analytical instrumentation utilized, in addition to consideration of other factors such as formulation cost.

Since only very small amounts of each candidate formulation are necessary for evaluation, the present invention makes it more feasible to test a very large number of candidate formulations, particularly because the combined total amount or volume of the formulations is manageable. Moreover, in the case of oligonucleotide markers, PCR amplification or alternative multiplication techniques permit identification of the oligonucleotide marker sequence where the amount recovered of an optimal formulation is too small for standard chemical analysis. Hence, the oligonucleotide marker is capable of being accurately identified with higher sensitivity than the formulation itself, which typically cannot be similarly multiplied or amplified. PCR amplification requires, however, that a PCR primer or other initiation sequence be incorporated into the oligonucleotide marker prior to association with a candidate formulation.

Oligonucleotides for use as markers in the present invention may be synthesized using any of the various techniques presently known in the art or subsequently developed. See, Gait, ed. (1984) Oligonucleotide Synthesis: A Practical Approach (IRL Press, Oxford) the disclosure of which is incorporated herein by reference. For example, oligonucleotides may be synthesized using the solid-phase method. Commonly used solid supports for this method of oligonucleotide synthesis include polystyrene, silica gel, glass beads, polyamide and cellulose paper. In a preferred

embodiment of the invention, the marker is a synthetic nucleic acid (DNA or RNA) oligonucleotide between about 12 and 100 nucleotides in length and having a sequence corresponding to a specific composition and component ratio of a candidate formulation. A variable region (or regions) contained within the oligonucleotide marker provides a sufficient number of oligonucleotide combinations to permit their use as unique markers to the candidate formulations. For example, using the four common nucleic acid bases for synthesis of the oligonucleotide marker yields at least 4^x possible unique combinations, where x is the number of bases in the variable region(s). Thus, where x=10, there are at least 4^{10}=1,048,576 unique oligonucleotide markers available. The length of the variable region(s) may be adjusted to provide an array of at least as many oligonucleotide combinations as there are candidate formulations to be evaluated. Association of the oligonucleotide marker with a candidate formulation may be obtained by encapsulation of the marker within or attachment of the marker to the individual assemblies or particles of the formulation.

Following synthesis, association of an oligonucleotide marker with the discrete assemblies of a given candidate formulation may be accomplished by a variety of methods. For example, in the case of liposome formulations, an oligonucleotide marker may be passively entrapped within the aqueous core of the liposome according to any of the various techniques known in the art for entrapping nucleic acids. Alternatively, an oligonucleotide marker may be actively loaded into preformed liposomes, according to methods disclosed in United States Patent No. 5,227,170, issued 7/13/93 entitled "Encapsulation Process," or they may be incorporated into lipid constructs and lipid-based carriers, according to methods disclosed in United States Patent Application Serial No. 08/434,465, filed 5/4/95 entitled "Nucleic Acid Ligand Complexes," both of which are incorporated herein by reference.

BNSDOCID: <WO_____9720076A1_I_>

Alternatively, an electrophoric marker may be associated with a candidate formulation. In the case of lipid-based multicomponent formulations, a lipid component may be chemically modified to incorporate the electrophoric marker directly by, for example, attachment through a linker. An example of the basic structure of an electrophoric marker-modified lipid is:

lipid-linker-OCH_2-$(CH_2)_n$-CH_2-O-Ar,

where n represents the number of carbon units in the variable-length alkyl chain and Ar is a haloarene moiety and may comprise, for example, C_6Cl_5, $C_6H_2Cl_3$ or $C_6H_2Cl_2F$. The lipid portion may be composed of any appropriate lipid compound, and is preferably, in the case of drug carriers for *in vivo* applications, pharmaceutically acceptable. Most preferably, the lipid is distearoyl phosphatidylethanolamine (DSPE). The linker can be any bifunctional agent which couples the lipid to the alkyl haloarene moiety. A set of resolvable markers may be generated by varying n and Ar. The value of n may, for example, range from 3 to 12.

The set of modified lipids is used to tag a candidate formulation by including only certain members of the set in a given candidate formulation at a level between about 0.0001 and 10% (most preferably between about 0.01 and 1%) by weight each. If, as in the marker set described above, there are potentially 30 members of the electrophoric marker set (i.e., n has 10 possible values, Ar may be one of three haloarene species), then accordingly, 2^{30} = 1,073,741,824 distinct formulations can be differentiated using this method. However, depending on the particular analytical technique employed, not all 30 members of the set may be distinctly resolvable. One skilled in the art though may readily determine a sufficiently large marker set by selecting those markers which can be adequately resolved by the given analytical technique. Smaller marker sets comprising fewer formulations may be utilized where evaluating fewer

candidate formulations. The concentration of the modified lipid for inclusion in a given formulation is governed on the lower end only by the detection limit of the identification technique; the upper concentration limit is set according to that amount which, when incorporated in a formulation, alters the physical or chemical properties of the formulation.

Various methods can be used to identify the electrophoric marker, including gas chromatography and mass spectrometry. Electron capture gas chromatography (ECGC) is particularly preferred for analysis of the electrophoric markers because it has excellent sensitivity and discrimination, allowing the electrophoric marker to be accurately identified at significantly lower levels than the formulation with which it is associated.

In a preferred embodiment, the electrophoric marker is attached to the component via a linker and the linker contains a photochemically labile bond such that the electrophoric marker is cleaved specifically upon irradiation and thereby released for facile analysis. One example of a photolabile moiety for inclusion in the linker is o-nitrobenzylcarbonate.

Following association of the markers (e.g., oligonucleotide or electrophoric) with their corresponding candidate formulations, the candidate formulations may, in the case of drug carrier optimization, be administered intravenously (i.v.) or by another administration route (e.g., topically, orally, subcutaneously) into an animal for evaluation with respect to expression of a desired set of properties. The candidate formulations may also be administered *in vitro* to cells in culture for evaluation.

In one embodiment of the invention, drug carrier formulations are optimized with respect to their ability to enter the cytoplasm of a cell, and further yet, the cell nucleus. For example, delivery of DNA to the nuclei of cells is necessary for effective gene therapy. Liposomes are useful as carriers for cellular drug delivery formulations utilizing therapeutic DNA and RNA oligonucleotides. Naked or otherwise unprotected

oligonucleotides in their phosphodiester form are quickly degraded in body fluids by intracellular and extracellular enzymes, such as exonucleases and endonucleases, before they reach the target cell. Encapsulation within liposomes protects oligonucleotides from enzymatic degradation and may increase cellular uptake and delivery as a result of phagocytosis of the liposomes. Cationic liposomes and lipid constructs comprising aminomannose-derivatized cholesterol are particularly effective for introducing nucleic acids into cells, as described in United States Patent Application Serial No. 08/386,577, filed 2/10/95 entitled "Lipid Constructs for Cytoplasmic Delivery of Agents," the disclosure of which is incorporated herein by reference. In the case of therapeutic nucleic acid delivery, the formulation typically comprises, in addition to the therapeutic nucleic acid, a distinct identification marker, which may itself be an oligonucleotide. In cases where delivery is being evaluated, the marker can serve as the surrogate drug as well.

For formula optimization for gene therapy applications, each candidate formulation optionally incorporates a nucleic acid sequence which is capable of producing expression of a signal in target cells in a characteristic and identifiable manner. For example, the formulations may be prepared to incorporate a plasmid which is known to encode for a protein that produces a fluorescent signal at a characteristic wavelength when irradiated. See, Chalfie *et al.* (1994) Science 263:802-805). Accordingly, detection of fluorescence attributable to such a protein indicates effective delivery of a given candidate formulation to the nuclei of the target cells, and consequently, identification of an optimal formulation for gene therapy applications.

Where oligonucleotide markers are used, short strands of DNA or RNA oligonucleotides (less than approximately 100 nucleotides in length) are synthesized with known sequences. In a preferred embodiment, each

oligonucleotide is flanked by defined 5' and 3' fixed regions that allow primer hybridization for Klenow extension, a primer annealing site for cDNA synthesis, PCR amplification and efficient T7 RNA polymerase transcription. Most preferably, the various oligonucleotide markers are then associated with the discrete assemblies of the candidate formulations and a database, cross-referencing the compositional data of each candidate formulation with the identity of its associated marker, is established. The marker may be associated with a candidate formulation by encapsulation within or attachment (by either physical or chemical means) to the discrete components of the candidate formulation. In the case of liposome formulations intended for *in vivo* evaluation, the oligonucleotide marker is preferably encapsulated within the aqueous interior for protection from degradative enzymes.

The candidate formulations, each tagged with a unique marker, are subsequently evaluated with respect to expression of a desired set of properties. For example, drug carriers incorporating oligonucleotide or electrophoric markers are injected i.v. into a tumor-containing animal and evaluated to determine the extent of drug carrier targeting to the tumor site. After a sufficient time, the tumor may be excised from the animal and analyzed for the presence of markers. Recovered oligonucleotide markers may be amplified using, e.g., PCR technology, and sequenced. Methods of sequencing DNA or RNA strands of this length are well known to those skilled in the art. Alternatively, electrophoric markers may be identified using mass spectrometry or gas chromatography (e.g., capillary GC or ECGC). The marker identity, once ascertained, is used to determine the corresponding initial formulation composition either directly (where the compositional data of a formulation is encoded within the marker) or by reference to a separate database. The formulation(s) recovered in the greatest amount is (are) typically considered optimal formulation(s). In

BNSDOCID: <WO_____9720076A1_I_>

some instances, however, optimal formulations may be those formulations which are recovered in the least amount following evaluation.

Where the number of candidate formulations for evaluation exceeds the practical limit for a single experiment, as determined by, for example, the total maximum volume which can be evaluated at one time, the process may be repeated successively to the extent necessary. One skilled in the art will readily appreciate the various embodiments of the present invention and considerable flexibility can be afforded the practical implementation of the claimed methods.

Where liposomes or lipid constructs constitute the formulation to be optimized, they may be constructed using a variety of amphiphilic molecules, which consist of polar (hydrophilic) and non-polar (lipophilic) portions. Examples of hydrophilic portions include phosphato, glycerylphosphato, carboxy, sulfato, amino, hydroxy, choline and other polar groups. Examples of non-polar groups are saturated or unsaturated hydrocarbons such as alkyl, acyl, alkenyl or other lipid groups. The chain length of each of these non-polar groups may additionally be varied to alter the physical and pharmacological properties of the resultant formulations. Also, sterols (e.g., cholesterol), surface-modifying groups (e.g., carbohydrates), and other pharmaceutically acceptable adjuvants, including anti-oxidants like alpha-tocopherol, may be included in the formulation. Other liposome or lipid-based drug carrier ingredients suitable for particular applications are known to workers skilled in the art. See, e.g., New, ed. (1992) Liposomes: A Practical Approach (IRL Press, Oxford). Similarly, where multicomponent formulations other than lipid-based assemblies are contemplated for optimization, appropriate components may be chosen by those skilled in the art.

Referring generally to FIG. 1, a flowchart summarizing one preferred embodiment of the present invention is depicted. This flowchart outlines

the method of the present invention for optimizing drug carrier formulations for *in vivo* targeting in a body. In this embodiment, an animal tumor cell is targeted using various drug carrier formulations.

Referring in detail to FIG. 1, solvents containing one or more components are conducted through a series of input tubes "A." The solvents from input channels "A" are mixed in specific proportions in chamber "B" to form a first candidate formulation, followed successively (in separate mixing steps) by the preparation of other candidate formulations. The candidate formulation in each given step (each formulation containing, for example, 100 μL) is then individually tagged with a unique marker (e.g., a synthetic oligonucleotide or an electrophoric moiety) supplied by "C." Each marker is programmed to correspond to a specific component ratio in each successive candidate formulation, such that each successive input stream from "C" comprises a solvent and one or more markers, synchronized with the discrete 100 μL zones of each candidate formulation. The procedure for associating a marker with a candidate formulation will depend on the nature of the association (e.g., encapsulation, external binding, inclusion as an integral component); appropriate procedures for associating markers with candidate formulations may be determined by one skilled in the art. In the case where the marker is an oligonucleotide encapsulated within liposomes, chamber "B" contains machinery necessary to prepare lipid vesicles (e.g., homogenizer). Once the formulations are tagged with the markers, it is unimportant whether the individual candidate formulations are maintained as discrete zone packets or are combined (i.e., admixed).

Optionally, the candidate formulations may be subjected to size selection before evaluation. For example, the candidate formulations may be passed through device "D" which separates and discharges components with sizes below a minimum cutoff value, and may then be conducted through filter "E" which passes components with sizes below a maximum

cutoff value.

The marked formulations may be eventually collected in vessel "F" and subsequently administered (e.g., by injection) into a test animal "G" having a tumor or tumors "H." Following an appropriate treatment time, a tumor is excised from the animal and is treated to release individual cells into solution. The released cells are then input into device "I" which separates cells from supernatant based on a signal associated with the cells. For example, "I" may comprise a fluorescence-activated cell sorter (FACS) which separates cells that express a fluorescence signal. The signal-expressing cells are then passed through "J," which analyzes and identifies the markers associated with the tumor cells (i.e., optimal formulation markers). In the case where the marker is an oligonucleotide, "J" may comprise instrumentation for amplifying the markers (e.g., by PCR) and sequencing the nucleic acids. Where the marker is an electrophoric haloarene, "J" may comprise, for example, mass spectrometric or gas chromatographic instrumentation. With the identity of the markers known, the initial compositions of the optimal formulations can be determined, for example, from the marker directly or by reference to database "K" which contains information cross-referencing the identity of a marker with the composition of the candidate formulation with which it was originally associated in chamber "B."

EXAMPLE 1

FORMULATION OPTIMIZATION FOR IN VIVO TARGETING OF LIPID-BASED DRUG CARRIERS USING OLIGONUCLEOTIDE MARKERS

Lipid-based drug carrier formulations are prepared by combining four component ingredients and optimized for their ability to target specific tissue sites *in vivo*.

Component samples consisting of the lipids dipalmitoylphosphatidylcholine (DPPC), dioleoylphosphatidylethanolamine (DOPE), dioleoyltrimethylammonia (DOTMA), and cholesterol are separately dissolved in ethanol to permit their individual introduction to mixing chamber "B" via input channels "A" shown in FIG. 1. DNA oligonucleotides (62-nucleotides in length) having the sequence:

5'- ATCCAATCTCTCCACATCTCTATACTATCATCACC - $N_1N_2N_3N_4N_5N_6N_7$-CTCACTCACTTCCATTCCAC -3'

(SEQ ID NO:1)

are prepared using solid phase synthesis. Synthesis of the variable base region N_1 through N_7 is programmed to provide a unique sequence for each oligonucleotide marker. The input channels "A" of FIG. 1 are programmed to increment each lipid channel in 10 mole% steps with respect to the other lipid channels which are also incremented in 10 mole% steps, taking into account the concentration of lipid in each solvent. In this example, the number of distinct oligonucleotide markers (using a variable string of seven bases) is $4^7 = 16{,}384$, sufficient to uniquely mark each possible candidate formulation.

Each oligonucleotide marker is associated with the discrete assemblies of its corresponding candidate formulation (volume = 100 μL) in chamber "B" of FIG. 1, where the candidate formulations are successively and uniquely tagged. The marked candidate formulations are filtered by size to remove free oligonucleotide marker and large lipid particles (e.g., aggregates). After collection of the size-selected formulations in vessel "F" of FIG. 1, the volume of the solution is reduced by membrane filtration (300 kilodalton cutoff) to provide a lipid concentration of 25 mg/mL.

The collected, concentrated mixture of marked candidate formulations is injected into the lateral tail vein of CD_2F_1 mice bearing P1798 lymphosarcoma tumors that had been implanted intradermally in the

BNSDOCID: <WO_____9720076A1_I_>

flanks of the mice 4 days prior to injection. An injection volume of 0.2 mL is used per mouse (each mouse approximately 20 grams weight) and a cohort of 10 mice is used. The mice are sacrificed 24 hours after injection and dissected. Blood, tumor, liver, spleen, kidney, heart, lungs, bone, intestines, brain, muscle and skin are separately collected. The tissues are immediately homogenized individually and a fraction enriched in the oligonucleotide marker is collected for each tissue. This fraction is used in a PCR amplification apparatus using primer sequences complementary to the 15 bases at the 5' and 3' ends of the encoding DNA whose sequence is provided above. The amplified DNA of each tissue is sequenced and decoded. Optimal formulations are identified by the extent of their accumulation in the tissues. From the oligonucleotide sequence, the optimal formulation may be identified by reference to the database containing information cross-referencing the identity of a particular oligonucleotide marker and the lipid formulation with which it was originally associated.

EXAMPLE 2

FORMULATION OPTIMIZATION FOR CELLULAR DELIVERY OF PLASMIDS USING ELECTROPHORIC MARKERS

Distearoylphosphatidylethanolamine (DSPE) is coupled by condensation to the electrophoric moiety shown in FIG. 2, and the product is purified by reverse phase chromatography. Two modified lipids are generated, where Ar=C_6Cl_5 and n=3 in one case and n=4 in the other. A spray-dried powder is then produced consisting of 79.9% (by weight) distearoylphosphatidylcholine (DSPC), 20.0% cholesterol (by weight), and 0.1% (by weight) of the n=3 modified DSPE. Hydration of 100 mg of the powder in 4.0 mL of water at 65°C is followed by sonication to yield unilamellar liposomes of 60 nm average diameter. A similar preparation is made using dioleoylphosphatidylethanolamine (DOPE)/6-aminomannose

cholesterol (1:1 mole ratio), with 0.1% n=4 marker.

A quantity of 1.0 mg plasmid DNA is added and incorporated into 1.0 mL of each of the two candidate liposome formulations. The plasmid encodes for a protein that produces a fluorescent signal upon irradiation. The formulation samples are applied simultaneously to a suspension of P1798 tumor cells in culture. The cells are incubated for 4 hours with the candidate liposome formulations and then sorted by fluorescence-activated cell sorting (FACS) using the fluorescent protein signal to gate the collection. The suspension of high fluorescence cells is centrifuged and the pellet is homogenized. The dried homogenate is irradiated at 365 nm to break the photolabile linker and release the electrophoric marker, which is then analyzed by ECGC. The predominant signal is from the n=4 marker, indicating the superiority of liposome formulations comprising DOPE/6-aminomannose cholesterol over similar liposome formulations comprising DSPC/cholesterol for delivery of plasmids into cells.

EXAMPLE 3

MULTIPLE FORMULATION ENCODING
USING ELECTROPHORIC MARKERS

The procedure and materials of Example 1 are followed except that electrophoric moieties rather than DNA oligonucleotides are used as markers.

The candidate formulations are tagged using fourteen programmed syringe pumps containing solutions of DSPE modified with electrophoric markers as in Example 2, comprising the series n=3 to n=12 for Ar=C_6H_5 and n=4 to n=7 for Ar=$C_6H_2Cl_3$. The number of distinct electrophoric markers is therefore 2^{14}=16,384. The syringe pumps are programmed to deliver 0.1% by weight of the modified lipids to mark each of the candidate formulations. The formation of lipid carriers and animal experimentation

are carried out as in Example 1.

After tissue collection, the samples are homogenized and dried. The samples are subjected to irradiation at 365 nm and the released electrophoric markers are analyzed by ECGC. The pattern of peaks is used to determine the formulation(s) most actively taken up in each tissue type.

Although the present invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that various changes and modifications may be made thereto without departing from the spirit or scope of the claims. In addition, changes in form and the substitution of equivalents are contemplated as circumstances may suggest or render expedient. Therefore, the foregoing description should not be construed to limit the scope of the present invention, which is set forth in the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: PAUL SCHMIDT

(ii) TITLE OF INVENTION: METHODS FOR OPTIMIZING MULTICOMPONENT FORMULATIONS

(iii) NUMBER OF SEQUENCES: 1

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Swanson & Bratschun, L.L.C.
(B) STREET: 8400 E. Prentice Ave., Suite 200
(C) CITY: Englewood
(D) STATE: Colorado
(E) COUNTRY: US
(F) ZIP: 80111

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WordPerfect 6.1

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Barry J. Swanson
(B) REGISTRATION NO.: 33,215
(C) REFERENCE NO.: NEX48/PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (303) 793-3333
(B) TELEFAX: (303) 793-3433

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATCCAATCTC TCCACATCTC TATACTATCA TCACCNNNNN NNCTCACTCA 50
CTTCCATTCC AC 62

CLAIMS:

1. A method for optimizing a multicomponent formulation comprising the steps of:

(a) preparing a plurality of candidate formulations of known initial composition;

(b) associating with each candidate formulation a marker of known identity, each marker being uniquely correlated to the candidate formulation with which it is associated;

(c) evaluating the marked candidate formulations for at least one desired property to determine an optimal formulation;

(d) identifying the marker which is associated with the optimal formulation; and

(e) determining the initial composition of the optimal formulation using the marker identity obtained in (d).

2. The method of claim 1 wherein the plurality of candidate formulations comprises a number greater than 10.

3. The method of claim 1 wherein said marker is selected from an oligonucleotide or an electrophoric moity.

4. The method of claim 3 wherein said oligonucleotide is selected from RNA or DNA.

5. The method of claim 3 wherein the oligonucleotide is between 12 and 100 nucleotides in length.

6. The method of claim 3 wherein the electrophoric marker is comprised of a haloarene moiety.

7. The method of claim 6 wherein said haloarene moiety is selected from the group consisting of C_6Cl_5, $C_6H_2Cl_3$ and $C_6H_2Cl_2F$.

8. The method of claim 6 wherein said electrophoric marker is comprised of a C_3 to C_{12} alkyl chain and a haloarene moiety.

9. The method of claim 8 wherein said electrophoric marker is attached to a lipid.

10. The method of claim 9 wherein said lipid is present at a concentration of about 0.0001% to 10% by weight of the formulation.

11. The method of claim 9 wherein said lipid is present at a concentration of about 0.01% to 1% by weight of the formulation.

12. The method of claim 9 wherein said electrophoric marker is attached to said lipid via a linker.

13 The method of claim 12 wherein said linker comprises a photolabile linkage.

14. The method of claim 1 wherein one or more of the formulation components comprises an amphiphilic compound.

15. The method of claim 14 wherein the amphiphilic compound is a lipid.

16. The method of claim 15 wherein the lipid is distearoylphosphatidylethanolamine.

BNSDOCID: <WO_____9720076A1_I_>

17. The method of claim 15 wherein the lipid is cationic.

18. The method of claim 1 wherein one or more of the formulation components comprises a sterol or a derivative thereof.

19. The method of claim 18 wherein the sterol is cholesterol.

20. The method of claim 18 wherein the sterol derivative is aminomannose-derivatized cholesterol.

21. The method of claim 1 wherein the candidate formulations comprise a drug carrier.

22. The method of claim 21 wherein the drug carrier is a liposome.

23. The method claim 1 whereby determination of the initial composition of the optimal formulation is accomplished by reference to a database which contains information cross-referencing the identity of the marker and the identity of its associated formulation.

24. The method of claim 1 wherein said plurality of candidate formulations are evaluated substantially simultaneously.

25. A method for optimizing a multicomponent formulation comprising the steps of:

(a) preparing a plurality of candidate formulations of known initial composition;

(b) associating with each candidate formulation a marker of known sequence;

(c) administering the marked candidate formulations to a collection of live cells;

(d) evaluating the marked candidate formulations for at least one desired property to determine an optimal formulation; and

(e) determining the initial composition of the optimal formulation.

26. The method of any one of claim 25 wherein the plurality of candidate formulations comprises a number greater than 10.

27. The method of claim 25 wherein said marker is an oligonucleotide.

28. The method of claim 26 wherein said marked formulations are evaluated by amplifying said oligonucleotide marker associated with the optimal formulation to an extent sufficient to permit sequencing of the marker.

29. The method of claim 28 wherein said oligonucleotide marker is amplified by a polymerase chain reaction.

30. The method of claim 25 wherein said marker is an electrophoric moity of known chemical formula.

31. The method of claim 28 wherein said electrophoric marker is evaluated using a gas chromatographic technique.

32. The method of claim 29 wherein said gas chromatographic technique is capillary gas chromatography.

33. The method of claim 29 wherein said gas chromatographic technique is electron capture gas chromatography.

34. The method of claim 30 wherein said electrophoric marker is evaluated using a mass spectrometric technique.

35. The method of claim 25 wherein the cell collection comprises an *in vitro* cell culture.

36. The method of claim 25 wherein the cell collection comprises a plant.

37. The method of claim 25 wherein the cell collection comprises a human or non-human animal body.

38. The method of claim 25 wherein one or more of the formulation components comprises an amphiphilic compound.

39. The method of claim 38 wherein the amphiphilic compound is a lipid.

40. The method of claim 39 wherein the lipid is distearoylphosphatidylethanolamine.

41. The method of claim 39 wherein the lipid is cationic.

42. The method of claim 25 wherein one or more of the formulation components comprises a sterol or a derivative thereof.

43. The method of claim 42 wherein the sterol is cholesterol.

44. The method of claim 43 wherein the sterol derivative is aminomannose-derivatized cholesterol.

45. The method of claims 25 wherein the candidate formulations comprise a drug carrier.

46. The method of claim 45 wherein the drug carrier is a liposome.

47. The method claim 25 whereby determination of the initial composition of the optimal formulation is accomplished by reference to a database which contains information cross-referencing the identity of the marker and the identity of its associated formulation.

48. The method of claim 25 wherein means for administration to the animal body is selected from the group consisting of intravenous, intramuscular, subcutaneous, percutaneous, oral, ingestion and topical application.

49. The method of claim 25 wherein the one or more desired properties comprises targeted delivery of the formulation to a specific tissue within a human or non-human animal body.

50. The method of claim 49 wherein the targeted tissue is a tumor.

51. The method of claim 25 wherein said plurality of candidate formulations are evaluated substantially simultaneously.

BNSDOCID: <WO_____9720076A1_I_>

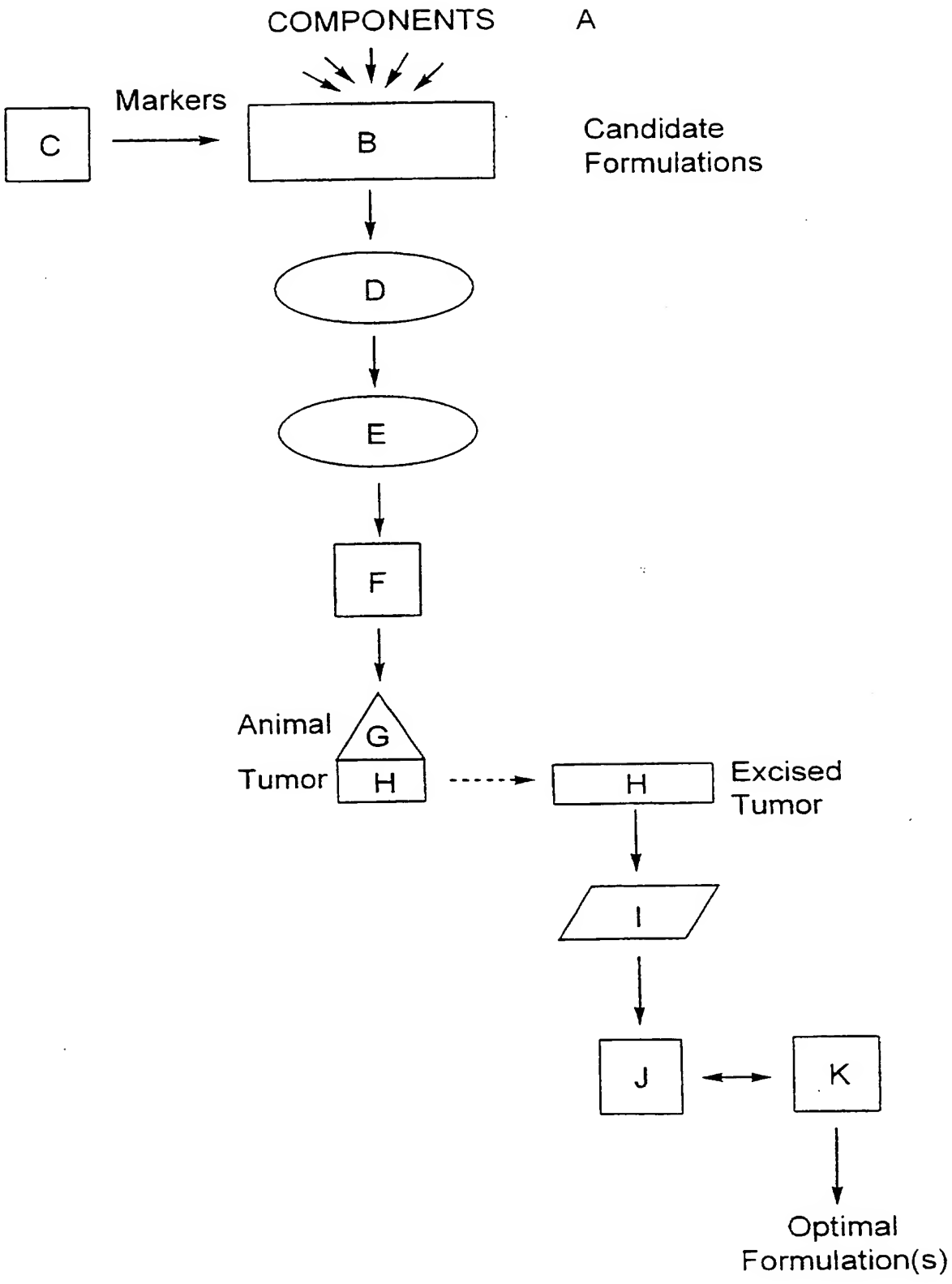
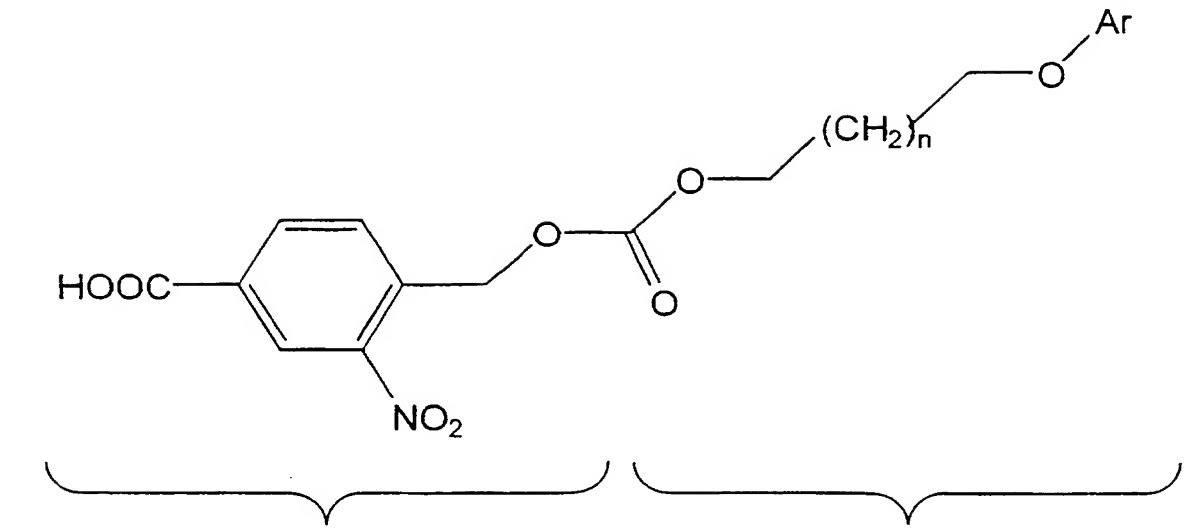

COMPONENTS
A
Markers
C
B
Candidate Formulations
D
E
F
Animal
Tumor
G
H
H
Excised Tumor
I
J
K
Optimal Formulation(s)

Fig. 1


Ar
O
(CH2)n
O
O
O
HOOC
NO2
Photolabile Linker
Electrophoric Marker

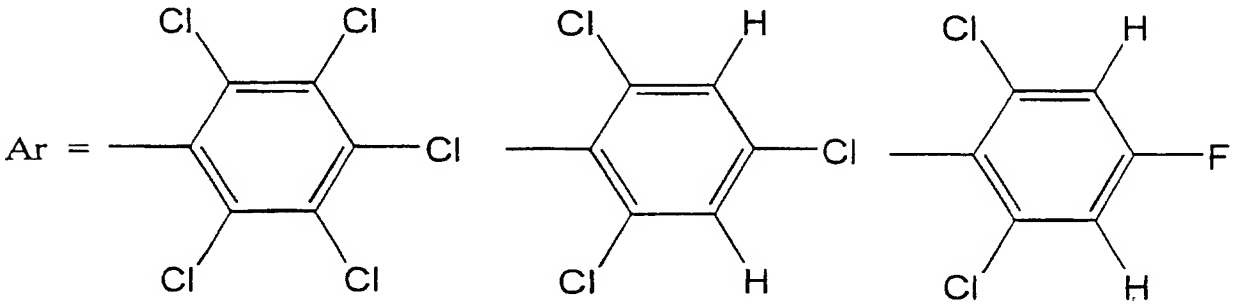

Ar =
Cl
Cl
Cl
Cl
Cl
Cl
Cl
H
Cl
H
Cl
H
Cl
H
F

Fig. 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/19030

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12Q 1/68; C12P 19/34
US CL :435/4, 6, 91.2; 436/94; 514/1, 44
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/4, 6, 91.2; 436/94; 514/1, 44

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, BIOSIS, CA, DERWENT
search terms: tag, marker, identifier, amplify, oligonucleotide, haloarene

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,451,661 A (WAN) 19 September 1995, column 3, lines 41-47 and column 9, line 51 to column 10, line 26.	9-51
X ---- Y	NEEDELS et al. Generation and screening of an oligonucleotide-encoded synthetic peptide library. Proc. Natl. Acad. Sci. USA. November 1993, Vol. 90, pages 10700-10704, especially pages 10703-10704.	1-5 ------- 14-29 and 35-51
X ----- Y	OHLMEYER et al. Complex synthetic chemical libraries indexed with molecular tags. Proc. Natl. Acad. Sci. USA. December 1993, Vol. 90, pages 10922-10926, especially pages 10923-10924.	1-3 and 6-8 ----- ---------- 9-26 and 30-51
Y	US 5,441,745 A (PRESANT et al.) 15 August 1995, column 5, lines 30-43	18-20 and 42-44

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 15 JANUARY 1997	Date of mailing of the international search report 11 FEB 1997
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer KENNETH R. HORLICK Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

BNSDOCID: <WO_____9720076A1_I_>

THIS PAGE BLANK (USPTO)